

Automated Event Service (AES): Efficient and Flexible Searching for Earth Science Phenomena

PI: Tom Clune, NASA GSFC

Objective

Develop an Automated Event Service system that:

- Methodically mines custom-defined events (e.g., tornadoes) in the reanalysis data sets of global atmospheric models.
- Enables researchers to specify their custom, numeric event criteria using a user-friendly web interface to search the reanalysis data sets.
- Supports Event Specification Language (ESL) for more flexibility and versatility.
- Contains a social component that enables the dynamic formation of collaboration groups for researchers to cooperate on event definitions of common interest.
- Provides rapid results via high performance computing and advanced search technologies.

Automated Event Services Utilize Big-Data Identify occurrences collaboration (events) of phenomena technologies to... Entities in the 4D Enable interactive and collaborative spatiotemporal space scientific data analysis on big data 9 Associate additional Share data and analysis methods interactive data and visual analytics with defining features ...in order to... extracted from data. Relieve scientists from data 4 Correlate defining features Big-Data technologies (SCIDB. POLARIS, HADOOP) management. of various phenomena in both space and time. Empower scientists to focus 5 Improve predictions of on science, and future events using correla Boost science productivity tions among phenomena for better decision making Big-Data Technology Science Vision Infrastructure Enablement

Approach

- Leverage advances in high-end computing and search technologies to create an efficient mechanism for searching reanalysis data for events.
- Build baseline system by custom integration of mature components: HPC cluster, MapReduce, Hadoop/Hbase.
- Develop ESL via analysis of representative events.
- Adapt advanced tree-based indexing strategies to efficiently support parameter-based event queries.
- Apply agile methodology: develop in small increments driven by use cases and synthetic tests.

Co-Is/Partners: Kwo-Sen Kuo, Caelum Research; Rahul Ramachandran, NASA MSFC

Key Milestones

 Import reanalysis data 	10/12
 Implement native indexing 	02/13
 Complete event web service 	05/13
 Complete basic web portal 	11/13
 Complete distributed event database 	02/14
 Design review for ESL 	05/14
 Demonstrate Blizzard use case on 30-node 	
commercial cloud cluster	07/14
 Complete multifaceted web portal 	11/14
 Complete tree-index search capability 	03/15
 Complete delivery of ESL 	03/15
 Testing and validation 	05/15

TRL_{in} = 2 TRL_{current} = 2

